GLOBAL WELL-POSEDNESS OF THE BENJAMIN-ONO EQUATION IN $H^1(\mathbf{R})$

TERENCE TAO

ABSTRACT. We show that the Benjamin-Ono equation is globally well-posed in $H^s(\mathbf{R})$ for $s \geq 1$. This is despite the presence of the derivative in the non-linearity, which causes the solution map to not be uniformly continuous in H^s for any s [15]. The main new ingredient is to perform a global gauge transformation which almost entirely eliminates this derivative.

1. INTRODUCTION

In this paper we study the Cauchy problem for the Benjamin-Ono equation

$$u_t + Hu_{xx} = uu_x; \quad u(0,x) = u_0(x), \tag{1}$$

where $u : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ is a real-valued function, H is the spatial Hilbert transform

$$Hu(x) := \frac{1}{\pi} p.v. \int_{\mathbf{R}} \frac{u(y)}{x-y} \, dy,$$

and u_0 lives in the inhomogeneous Sobolev space H_x^s for some $s \in \mathbf{R}$. Recall that these spaces can be defined via the spatial Fourier transform

$$\hat{f}(\xi) := \int_{\mathbf{R}} e^{-2\pi i x \xi} f(x) \, dx$$

by the formula

$$\|f\|_{H_x^s} := \left(\int_{\mathbf{R}} \langle \xi \rangle^{2s} |\hat{f}(\xi)|^2 \, d\xi \right)^{1/2},$$

where $\langle \xi \rangle := (1 + |\xi|^2)^{1/2}$.

This equation is a model for one-dimensional waves in deep water [2], and is completely integrable [1]. The Cauchy problem for this equation has been extensively studied [8], [9], [11], [14], [15], [23], [24], [29]. It is known that this equation is globally well-posed in the Sobolev space H_x^s if s is sufficiently large; for instance it is known that if $u_0 \in H_x^3$ (for instance), then there exists a unique (classical) global solution u to (1) which lies in the space $C_t^0 H_x^3$; we then define the solution operator $S(t) : H_x^3 \to H_x^3$. Furthermore, it is known that for every fixed t, the map $S(t) : H_x^3 \to H_x^3$ is continuous (see e.g. [24]), although it is not analytic or even uniformly continuous on bounded sets (see [15], [20]).

1991 *Mathematics Subject Classification.* 35J10.

This paper is concerned with the question of whether S can be extended to rougher classes of initial data. The above H_x^3 result has already been improved by several authors; for instance, Iorio [11] obtained local well-posedness in H^s for $s > 3/2$, which was then improved by Ponce [23] to well-posedness in $s \geq 3/2$. More recently, Koch and Tzvetkov [14] have improved this to $s > 5/4$.

The Hilbert transform behaves in some ways like the constant $\pm i$; for instance, we have that H is anti-self-adjoint with $H^2 = -1$. A more precise relationship is given by the well-known formula

$$\widehat{Hf}(\xi) = -i\mathrm{sgn}(\xi)\hat{f}(\xi) \tag{2}$$

which can easily be verified by contour integration. Thus, the Benjamin-Ono equation (1) is heuristically of the form

$$u_t \pm iu_{xx} = uu_x.$$

We remark that quadratic non-linear Schrödinger equations of the form

$$u_t \pm iu_{xx} = |u|^2 \tag{3}$$

have been studied in [17], and local well-posedness was established for such equations in $H^s(\mathbf{R})$ for[1] $s > -1/4$. Note that it is not *a priori* obvious that one can make sense (as a distribution) of non-linearities such as $|u|^2$ for solutions in negative Sobolev spaces, but it turns out that the Schrödinger equation on the line has sufficient smoothing properties to make this possible.

There is thus a substantial gap - more than a derivative - in regularity between the previous local well-posedness theory for the Benjamin-Ono equation and for the quadratic NLS equation. We believe that this gap is mostly artificial, in that the well-posedness theory for the Benjamin-Ono equation can be lowered much further, to become closer to that for the quadratic NLS equation. As a first step in this direction we give

Theorem 1.1. *Let $R > 0$ and $s_0 \geq 1$. Then there exists a time $T = T(R, s_0) > 0$ such that for every time $-T \leq t \leq T$, the solution operator $S(t) : H_x^3 \to H_x^3$ can be continuously and uniquely extended to the ball*

$$B(0, R) := \{u_0 \in H_x^1 : \|u_0\|_{H_x^1} \leq R\},$$

endowed with the H^1 topology, and $S(t)u_0$ solves (1) in the sense of distributions. In addition, if $u_0 \in B(0, R) \cap H_x^{s_0}$, then $S(t)u_0 \in C_{[-T,T]}^0 H_x^{s_0}$, and indeed

$$\|S(t)u_0\|_{C_{[-T,T]}^0 H_x^{s_0}} \leq C(s_0, R)\|u_0\|_{H_x^{s_0}}. \tag{4}$$

Here of course we give $C_{[-T,T]}^0$ the sup norm topology:

$$\|S(t)u_0\|_{C_{[-T,T]}^0 H_x^{s_0}} := \sup_{t \in [-T,T]} \|S(t)u_0\|_{H_x^{s_0}}.$$

Indeed, the solution map is continuous from $H_x^{s_0} \cap B(0, R)$, with the $H_x^{s_0}$ topology, to $C_{[-T,T]}^0 H_x^{s_0}$.

---

[1]It is not known whether this value of s is sharp, however the bilinear estimate used in [17] to establish this result certainly fails [22] for $s < -1/4$, and indeed can be used to show that the solution map is not analytically well-posed. See also [21].

Finally, the solution map is Lipschitz as a map from $B(0, R)$, with the L_x^2 topology(!), to $C_{[-T,T]}^0 L_x^2$, or in other words that

$$\|S(t)u_0 - S(t)\tilde{u}_0\|_{L_x^2} \leq C(s_0, R)\|u_0 - \tilde{u}_0\|_{L_x^2} \tag{5}$$

for all $t \in [-T, T]$ and $u_0, \tilde{u}_0 \in B(0, R)$.

This improves[2] on the earlier result of Koch and Tzvetkov [14], which obtained local well-posedness in H_x^s for $s > 5/4$. Also, in [15] it is remarked that the solution map is not uniformly continuous from H_x^s to H_x^s for any s, because of the derivative in the non-linearity and the relatively weak smoothing effects of the linear part of the equation, and so a direct iteration map cannot work; note however that the construction in [15] does not prohibit the solution map being uniformly continuous or Lipschitz in a weaker topology such as L_x^2. We circumvent this lack of uniform continuity by applying a gauge transform (a variant of the classical Cole-Hopf transformation) to effectively remove the derivative (or at least the worst terms which involve[3] the derivative) from the non-linearity. This gauge transform may well be related to the complete integrability of the Benjamin-Ono equation (see e.g. [1]), although we do not know of a direct connection between our methods and the theory of integrable systems.

Our methods used are fairly elementary, relying mostly on the algebraic manipulations of the gauge transform, as well as standard Strichartz estimates for the Schrödinger equation. Littlewood-Paley projections play a small role (except in obtaining the H^1 continuity estimate), but mostly we need them to separate the action of the Hilbert transform on the positive and negative frequencies. The Lipschitz bound (5) in particular is very simple and relies on the energy method combined with the $L_t^4 L_x^\infty$ Strichartz estimate.

For smooth (e.g. H_x^3) solutions, it is easy to check that the L_x^2 norm $\int u^2$ is an invariant of the flow (1):

$$\int_{\mathbf{R}} u(t, x)^2 \, dx = \int_{\mathbf{R}} u_0(x)^2 \, dx. \tag{6}$$

Somewhat less obvious conserved quantities include the Hamiltonian

$$\int_{\mathbf{R}} uHu_x - \frac{1}{3}u^3 \, dx, \tag{7}$$

as well as the H^1-type quantity

$$\int_{\mathbf{R}} u_x^2 - \frac{3}{4}u^2 Hu_x - \frac{1}{8}u^4 \, dx; \tag{8}$$

in fact the Benjamin-Ono equation is completely integrable (see e.g. [1]) and has an infinite number of such conserved quantities. From the above conservation laws

---

[2] Since the submission of this paper, the author has learned (C. Kenig, personal communication) that a similar improvement by different methods has been achieved by C. Kenig and K. Koenig, for $s > 9/8$.

[3] Specifically, we can eliminate the terms involving an interaction of very low and very high frequencies, where the derivative falls on the very high frequency; this interaction was specifically identified as the culprit to lack of uniform continuity in [15].

and the Gagliardo-Nirenberg inequality it is easy to verify the *a priori* bound

$$\|u(t)\|_{H_x^1} \leq C(\|u_0\|_{H_x^1}) \tag{9}$$

for all $t \in \mathbf{R}$. *A priori* this bound can only be proven for smooth solutions, but by the continuity properties obtained in Theorem 1.1 we see that they in fact hold for H_x^1 solutions as long as the solution exists. From this and iterating Theorem 1.1, we thus obtain

Corollary 1.2. *The equation* (1) *is globally well-posed in H_x^s for all $s \geq 1$.*

Interestingly, a global existence result has recently been obtained for the viscous Benjamin-Ono equation in [6]; it thus seems feasible to show that the solutions to the viscous Benjamin-Ono equation converges to the inviscid Benjamin-Ono equation in the zero viscosity limit, although we do not pursue this question here.

The author is indebted to Nikolay Tzvetkov and Jim Colliander for very helpful discussions on this problem, and to Felipe Linares for some help with the references. The author also thanks Martin Hadac for pointing out an error in the exposition in a previous version of the paper.

2. Notation

Fix $s_0 \geq 1$. We use the notation $A \lesssim B$ or $A = O(B)$ to denote the estimate that $|A| \leq C_{s_0} B$, where C_{s_0} is a quantity depending on s_0 but not on ε. If X is a Banach space, we use $O_X(B)$ to denote any element in X with norm $O(B)$. We use $\langle x \rangle$ to denote the quantity $\langle x \rangle := (1 + |x|^2)^{1/2}$.

We observe the Riesz decomposition

$$1 = P_- + P_+$$

where $P_\pm$ are the Fourier projections to $\pm[0, \infty)$; from (2) we observe that

$$iH = P_- - P_+. \tag{10}$$

Let ψ be a bump function adapted to $[-2, 2]$ and equal to 1 on $[-1, 1]$. We define the Littlewood-Paley operators P_k and $P_{\leq k} = P_{<k+1}$ for $k \geq 0$ by defining

$$\widehat{P_{\leq k} f}(\xi) := \psi(\xi/2^k)\hat{f}(\xi)$$

for all $k \geq 0$, and $P_k := P_{\leq k} - P_{\leq k-1}$ (with the convention $P_{\leq -1} = 0$). Note that all the operators P_k, $P_{\leq k}$ are bounded on all translation-invariant Banach spaces, thanks to Minkowski's inequality. We define $P_{>k} := P_{\geq k-1} := 1 - P_{\leq k}$.

From Plancherel we have the bound

$$\|f\|_{H_x^s} \sim \Big(\sum_{k=0}^{\infty} \|P_k f\|_{H_x^s}^2\Big)^{1/2} \sim \Big(\sum_{k=0}^{\infty} 2^{ks} \|P_k f\|_{L_x^2}^2\Big)^{1/2} \tag{11}$$

for any $s \in \mathbf{R}$.

We define $P_{lo} := P_0$, $P_{hi} := P_{>0}$, and $P_{\pm hi} := P_{\pm}P_{hi}$; thus $1 = P_{lo} + P_{+hi} + P_{-hi}$. Observe that if F is real, then $P_{lo}F$ is also real; this will be important when we compute expressions such as $e^{iP_{lo}F}$. We also define $P_{LO} := P_{\leq 2}$, $P_{HI} := P_{>2}$, and $P_{\pm HI} := P_{\pm}P_{HI}$, so $1 = P_{LO} + P_{+HI} + P_{-HI}$.

3. THE GAUGE TRANSFORMATION, AND A PRIORI STRICHARTZ ESTIMATES

We now begin the proof of Theorem 1.1. Fix s_0. We first observe from the scale invariance

$$u(t,x) \mapsto \frac{1}{\lambda}u\left(\frac{t}{\lambda^2}, \frac{x}{\lambda}\right)$$

of the equation (1), that we may rescale the H^1 norm R to be small. In particular we may assume that $R \ll \varepsilon^2$, and we will now obtain local well-posedness on the time[4] interval $[-1, 1]$ (i.e. we set $T := 1$).

From the Strichartz theory of the one-dimensional Schrödinger equation (see e.g. [12]) we have the global Strichartz estimate

$$\|u\|_{L_t^4 L_x^\infty} + \|u\|_{L_t^\infty L_x^2} \lesssim \|u(0)\|_{L_x^2} + \|u_t \pm iu_{xx}\|_{L_t^1 L_x^2}$$

for any spacetime test function $u(t,x)$ and either choice of sign $\pm$. Applying this to $P_{\pm}u$ and summing, we obtain in particular that

$$\|u\|_{L_t^4 L_x^\infty} + \|u\|_{L_t^\infty L_x^2} \lesssim \|u(0)\|_{L_x^2} + \|u_t + Hu_{xx}\|_{L_t^1 L_x^2},$$

and similarly for P_-u and P_+u. Differentiating this, we obtain

$$\|u\|_{S^k} \lesssim \|u(0)\|_{H_x^k} + \|u_t + Hu_{xx}\|_{L_t^1 H_x^k} \tag{12}$$

for any integer $k \geq 0$, where S^k denotes the Strichartz norm

$$\|u\|_{S^k} := \|u\|_{L_t^4 C_x^k} + \|u\|_{L_t^\infty H_x^k}.$$

In light of this, it is reasonable to expect that H_x^1 solutions of the Benjamin-Ono equation are bounded in S^1. We shall now show that this is in fact the case, at least for smooth solutions.

Theorem 3.1. *Let u be an H_x^3 solution to (1) with $u(0) = O_{H^1}(\varepsilon^2)$. Then we have*

$$u = O_{S^1([-1,1]\times\mathbf{R})}(\varepsilon^2). \tag{13}$$

Remark. From the classical global well-posedness theory in H_x^3 ([24]; see also [11], [23], [14]) we then know that as soon as u_0 is in H_x^3, then the solution u is in $C_t^0 H_x^3$ and exists globally in time. Note however that our bounds will be independent of the H_x^3 norm of u_0, and once we obtain some continuity estimates we will be able to pass to the limit and obtain a version of this theorem for general H_x^1 solutions.

---

[4]In principle, this scaling should yield a time of existence $T \sim \|u\|_{H^1}^{-4/3}$ when the data is large, since the rescaling shrinks the $\dot{H}^1$ norm by a factor of $O(\lambda^{-3/2})$. However our argument does not quite achieve this because the L_x^2 component of the H_x^1 norm does not scale so well, by a factor of only $O(\lambda^{-1/2})$. It is quite likely one can modify this argument to fix this problem, although for our purposes this is unnecessary because the H^1 conservation law yields a global existence result anyway.

Remark. The $L_t^\infty H_x^1$ component of the Strichartz norm S^1 can of course be controlled using the conserved quantities of the Benjamin-Ono equation, and in particular (8). However, we will refrain from using this conservation law here as it does not adapt well to frequency envelopes, which we will introduce in later sections. We will however take advantage of the L^2 conservation law (6) in order to easily deal with the low frequency portions of the solution (although it is possible to treat those without recourse to any conservation law).

Proof Henceforth we restrict all spacetime norms to the slab $[-1,1] \times \mathbf{R}$. A standard iteration method will not work here, because the linear part of the Benjamin-Ono equation does not have enough smoothing to compensate for the derivative in the non-linearity. To resolve this, we will gauge transform the equation (1) into a more manageable form, where the derivative term has been moved onto a low frequency factor (cf. the gauge transformations for wave maps in [25], [26], [13], [28]).

The first step is to take an antiderivative of (1), introducing a spatial primitive $F(t,x)$ of $u(t,x)$. We first construct $F(t,0)$ on the time axis $x = 0$ by solving the ODE

$$\partial_t F(t,0) + \frac{1}{2} H u_x(t,0) = \frac{1}{4} u(t,0)^2; \quad F(0,0) := 0$$

and then constructing $F(t,x)$ for general x by the ODE

$$\partial_x F(t,x) = \frac{1}{2} u(t,x). \tag{14}$$

Note that $F(t,x)$ is necessarily real-valued. From (1) we see that

$$\partial_x(F_t + H F_{xx}) = \partial_x(F_x^2)$$

while from our construction on the time axis we know that

$$F_t(t,0) + H F_{xx}(t,0) = F_x^2(t,0).$$

Thus we have the equation

$$F_t + H F_{xx} = F_x^2. \tag{15}$$

holding globally in spacetime.

Remark. If we replaced H by $-i$ (as one could do if F were somehow restricted to positive frequencies) then one could solve this equation $F_t - i F_{xx} = F_x^2$ explicitly using the Cole-Hopf transformation[5] $w := e^{-iF}$, since one can then easily verify that w solves the free Schrödinger equation $w_t - i w_{xx} = 0$.

Note that while F is quite smooth, being a primitive of an H_x^3 function, it does not necessarily decay at spatial infinity $|x| \to \infty$; in fact, it is not even bounded, since H_x^3 functions need not be absolutely integrable. However, this problem is purely in the low frequencies of F; it is clear that $P_{hi}F$ is bounded (and in H_x^4), since ∂_x is essentially an isometry from H_x^4 to H_x^3 on these frequency ranges.

---

[5]Actually, this is a complexified version of the classical Cole-Hopf transformation [4], [10], used to solve the inviscid Burger's equation. A similar transformation was also used by Nirenberg (see e.g. [18]) to solve the scalar non-linear wave equation $F_{tt} - \Delta F = |F_t|^2 - |\nabla F|^2$. We thank Jim Colliander for pointing out the relevance of the Cole-Hopf transformation to this problem.

The function e^{-iF} is clearly bounded, but obviously does not decay at infinity. However, its derivative $\partial_x e^{-iF} = -iF_x e^{-iF} = -\frac{i}{2} u e^{-iF}$ can easily be shown to be in H_x^3. In particular we can meaningfully talk about Fourier projections of e^{-iF} as long as we stay away from the frequency origin; for instance, $P_{hi} e^{-iF}$ is well defined. We can also define projections such as $P_{lo} e^{-iF}$ by defining $P_{lo} e^{-iF} := e^{-iF} - P_{hi} e^{-iF}$. Similar considerations apply of course to F itself.

From (14) it is clear that in order to get bounds on u it will suffice to obtain appropriate bounds on F which are of one higher degree of regularity. Inspired by the Cole-Hopf transformation mentioned earlier, we introduce the complex-valued field $w : \mathbf{R} \times \mathbf{R} \to \mathbf{C}$ defined by

$$w := P_{+hi}(e^{-iF}); \tag{16}$$

note that this projection is well defined since e^{-iF} has a derivative in H_x^3. If we then expand out $w_t - iw_{xx}$ and apply (15) and (10), we obtain

$$
\begin{aligned}
w_t + H w_{xx} &= w_t - i w_{xx} \\
&= P_{+hi}((-iF_t - F_{xx} + iF_x^2)e^{-iF}) \\
&= P_{+hi}((iHF_{xx} - F_{xx})e^{-iF}) \\
&= -2P_{+hi}(P_-(F_{xx})e^{-iF}) \\
&= -2P_{+hi}(P_-(F_{xx})w) - 2P_{+hi}(P_-(F_{xx})P_{lo}(e^{-iF}));
\end{aligned}
\tag{17}
$$

the last identity follows from (16) and the fact that $P_{+hi}(P_-(F_{xx})P_{-hi}(e^{-iF}))$ vanishes. Of the two terms, the second term is much better than the first, indeed it, and both of its factors, must necessarily be very low frequency. The more interesting term is the first one.

This gives us an equation for w in terms of F. Note that the very worst type of interaction in the Benjamin-Ono equation, in which a very low frequency is interacted with a very high frequency, is almost completely absent in this equation, since both derivatives will fall on the low frequency. (The F_{xx} term must have lower frequency than w, otherwise $P_-(F_{xx})w$ will have negative frequency and thus vanish when P_{+hi} is applied.)

Now that we have a good equation for w, we must go back and obtain a good equation for F. Define $F_{LO} := P_{LO}F$, $F_{+HI} := P_{+HI}F$, and $F_{-HI} := P_{-HI}F$, thus $F = F_{LO} + F_{+HI} + F_{-HI}$. Since F is real, we observe that F_{LO} is also real, and that $F_{-HI} = \overline{F_{+HI}}$. We differentiate (16) and compute

$$
\begin{aligned}
w_x &= -iP_{+hi}(F_x e^{-iF}) \\
&= -iP_{+hi}((F_{+HI})_x e^{-iF}) - iP_{+hi}((F_{LO} + F_{-HI})_x e^{-iF}) \\
&= -i(F_{+HI})_x e^{-iF} + i(P_{lo} + P_{-hi})((F_{+HI})_x e^{-iF}) - iP_{+hi}((F_{LO} + F_{-HI})_x e^{-iF})
\end{aligned}
$$

and hence, multiplying by ie^{iF} and rearranging, we obtain an equation

$$\partial_x F_{+HI} = ie^{iF} w_x + e^{iF} E \tag{18}$$

for F_{+HI}, where the error term E is given explicitly by

$$E := (P_{lo} + P_{-hi})(e^{-iF}(F_{+HI})_x) - P_{+hi}((F_{-HI})_x e^{-iF})P_{+hi}((F_{LO})_x e^{-iF}).$$
(19)

One can think of the formula (18) as a version of the Bony linearization formula for e^{iF}.

To prove (13) it will suffice by (14) to prove the estimates

$$\partial_x P_{\pm} F = O_{S^1}(\varepsilon^2); \quad w = O_{S^2}(\varepsilon^2).$$
(20)

By a standard continuity argument using the H_x^3 global well-posedness theory, it will suffice to prove this estimate assuming the bootstrap hypothesis

$$\partial_x P_{\pm} F = O_{S^1}(\varepsilon); \quad w = O_{S^2}(\varepsilon).$$
(21)

We now deduce (20) from (21). We first deal with the low frequencies F_{LO}, which are easy to handle. From the L^2 conservation law (6) and the smallness of the initial data in H_x^1 (and hence in L_x^2) we clearly have $u = O_{L_t^\infty L_x^2}(\varepsilon^2)$; applying P_{LO} and then Bernstein's inequality (and Hölder in time) we thus see that

$$\partial_x P_{\pm} F_{LO} = O_{S^1}(\varepsilon^2)$$
(22)

as desired.

Now we verify (20) for w. We begin with the initial data $w(0)$. From (14) and the hypothesis on u, we have $F_x(0) = O_{H_x^1}(\varepsilon^2)$, and so from the chain rule and Sobolev (and the fact that $e^{-iF(0)}$ is bounded, since $F(0)$ is real) it is easy to see that $\partial_x e^{-iF(0)} = O_{H_x^1}(\varepsilon^2)$. From (16) we thus have

$$w(0) = O_{H_x^2}(\varepsilon^2).$$

In order to verify (20) for w, it thus suffices by (12) to verify that

$$w_t + H w_{xx} = O_{L_t^1 H_x^2}(\varepsilon^2).$$

By (17) it thus suffices to show that

$$P_{+hi}(P_-(F_{xx})w), P_{+hi}(P_-(F_{xx})P_{lo}(e^{-iF})) = O_{L_t^1 H_x^2}(\varepsilon^2).$$

We first consider the term $P_{+hi}(P_-(F_{xx})w)$. From (21) we see that $P_-(F_{xx}) = O_{L_t^4 L_x^\infty}(\varepsilon)$ and $w = O_{L_t^\infty H_x^2}(\varepsilon)$. So the claim will follow from a Hölder in time, and the following paraproduct estimate.

Lemma 3.2. *For any functions $f(x)$ and $g(x)$, we have*

$$\|P_{+hi}(P_-(f)g)\|_{H_x^2} \lesssim \|P_-(f)\|_{L_x^\infty} \|g\|_{H_x^2}.$$

Proof For any integer $k \geq 0$, observe that

$$P_+ P_k(P_-(f)g) = P_+ P_k(P_-(f)P_{\geq k-1}g)$$

and hence (by the L^2 boundedness of $P_+ P_k$ and Hölder)

$$\|P_+ P_k(P_-(f)g)\|_{L_x^2} \lesssim \|P_-(f)\|_{L_x^\infty} \|P_{\geq k-1}g\|_{L_x^2}.$$

Multiplying by 2^{2k}, and then square summing in k, we obtain

$$\|P_{+hi}(P_-(f)g)\|_{H_x^2} \lesssim \|P_-(f)\|_{L_x^\infty}(\sum_k(2^{2k}\|P_{\geq k-1}g\|_{L_x^2})^2)^{1/2}.$$

But by Plancherel we readily see that the latter factor on the right-hand side is $O(\|g\|_{H_x^2})$, and the claim follows. $\blacksquare$

Now we estimate $P_{+hi}(P_-(F_{xx})P_{lo}(e^{-iF}))$. Observe from all the frequency localizations that we may freely replace F_{xx} by $P_{lo}F_{xx}$. But from (22) and frequency localization we see that $P_-P_{lo}F_{xx} = O_{L_t^\infty H_x^2}(\varepsilon^2)$, while the boundedness of e^{-iF} and frequency localization we have $P_{lo}e^{-iF} = O_{L_t^\infty C_x^2}(1)$. Multiplying these two estimates we see that $P_{+hi}(P_-(F_{xx})P_{lo}(e^{-iF})) = O_{L_t^1 H_x^2}(\varepsilon^2)$ as desired.

Finally, we verify (20) for F_{HI}. By conjugation invariance of the S^k spaces it suffices to do this for F_{+HI}, since $F_{-HI} = \overline{F_{+HI}}$. By (18) it thus suffices to show that

$$e^{iF}w_x, e^{iF}E = O_{S^1}(\varepsilon^2).$$

From (21) and Sobolev we know that $F_x = O_{L_t^\infty H_x^1}(\varepsilon) = O_{L_t^\infty C_x^0}(\varepsilon)$, and hence $e^{\pm iF} = O_{L_t^\infty C_x^1}(1)$. In particular, we see that multiplication by $e^{\pm iF}$ is a bounded operation on S^1. Thus it will suffice to show that

$$w_x, E = O_{S^1}(\varepsilon^2).$$

The claim for w_x has already been proven, so it suffices to estimate E. We expand E using (19). Consider the first term $(P_{lo} + P_{-hi})(e^{-iF}(F_{+HI})_x)$ in (19). Note that we may freely replace e^{-iF} by $P_{-hi}e^{-iF}$ due to all the frequency projections. From (21) we have $F_x = O_{L_t^\infty H_x^1}(\varepsilon)$, and hence $\partial_x e^{-iF} = -iF_x e^{-iF} = O_{L_t^\infty H_x^1}(\varepsilon)$. Hence $P_{-hi}e^{-iF} = O_{L_t^\infty H_x^2}(\varepsilon) = O_{L_t^\infty C_x^1}(\varepsilon)$ by Sobolev. Since $(F_{+HI})_x = O_{S^1}(\varepsilon)$, we thus easily see from Leibnitz and Hölder that

$$(P_{lo} + P_{-hi})(e^{-iF}(F_{+HI})_x) = O_{S^1}(\varepsilon^2)$$

as desired (note that P_{lo} and P_{-hi} are bounded on S^1).

The second term $P_{+hi}((F_{-HI})_x e^{-iF})$ of (19) is treated similarly, so we turn now to $P_{+hi}((F_{LO})_x e^{-iF})$. From (22) we have that $(F_{LO})_x = O_{S^1}(\varepsilon^2)$; since multiplication by e^{-iF} is bounded on S^1, we see that $P_{+hi}((F_{LO})_x e^{-iF}) = O_{S^1}(\varepsilon^2)$ as desired. This completes the proof of Theorem 3.1. $\blacksquare$

4. THE LIPSCHITZ BOUND

The objective of this section is to show that the solution map is Lipschitz continuous from $H_x^3 \cap B(0, R)$, in the L^2 topology, to $C_{[-T,T]}^0 L_x^2$. More precisely:

Theorem 4.1. Let $0 < R \ll \varepsilon^2$. Then (5) holds for all $u_0, \tilde{u}_0 \in H_x^3 \cap B(0, R)$.

Proof This will be an easy energy method argument. Write $u(t) = S(t)u_0$ and $\tilde{u}(t) = S(t)\tilde{u}_0$, and let $v(t) := \tilde{u}(t) - u(t)$ denote the difference between the two

solutions. Then v obeys the difference equation

$$v_t + Hv_{xx} = \frac{1}{2}(\tilde{u}^2 - u^2)_x = (Uv)_x$$

where $U := \frac{\tilde{u}+u}{2}$. Now we compute the growth of the L^2 norm of v:

$$\partial_t \int v(t)^2 \, dx = 2 \int v(t)v_t(t) \, dx$$
$$= 2 \int v(t)(v_t(t) + Hv_{xx}(t)) \, dx$$
$$= 2 \int v(t)(U(t)v(t))_x \, dx$$
$$= -2 \int v_x(t)U(t)v(t) \, dx$$
$$= - \int (v^2(t))_x U(t) \, dx$$
$$= \int v^2(t)U_x(t) \, dx$$
$$= O(\|U_x(t)\|_{L_x^\infty}) \int v(t)^2 \, dx.$$

By Gronwall's inequality, we thus have

$$\sup_{t \in [-1,1]} \int v(t)^2 \, dx \le \exp(\|U_x\|_{L_t^1 L_x^\infty}) \int v(0)^2 \, dx.$$

However, from Theorem 3.1 we have $U_x = O_{L_t^4 L_x^\infty}(\varepsilon^2) = O_{L_t^1 L_x^\infty}(\varepsilon^2)$ and hence

$$\sup_{t \in [-1,1]} \int v(t)^2 \, dx \lesssim \int v(0)^2 \, dx.$$

The claim (5) then follows.   ■

From Theorem 4.1 we know that the solution map $S(t)$ is a Lipschitz map from $H_x^3 \cap B(0,R)$, with the L_x^2 topology, to $C_t^0 H_x^3([-1,1] \times \mathbf{R})$, with the L_x^2 topology. Of course, in the L_x^2 topology, $H_x^3 \cap B(0,R)$ is dense in $B(0,R)$. Thus we may uniquely complete the solution map to be a Lipschitz continuous map from $B(0,R)$ with the H_x^1 topology, to $C_t^0 H_x^1$ with the H_x^1 topology; this of course proves (5) for all choices of initial data in $B(0,R)$. Also we see from (5) that the solution operator $S(t)$ constructed this way is unique in the class of limits of smooth solutions.

However, we are still missing several components of Theorem 1.1. Firstly, we have not yet shown that $S(t)$ is continuous in the H_x^1 topology (as opposed to the L_x^2 topology). Also, we have not yet shown that the solution operator constructed by the above limiting procedure is a solution in the sense of distributions. Finally, we have not shown the persistence of regularity bound (4). To obtain all of these additional properties of the solution map $S(t)$, we will revisit the a priori bound in Theorem 3.1, but this time we will introduce the machinery of *frequency envelopes*, as used recently in wave maps in [25], [26], [28].

5. FREQUENCY ENVELOPES

In order to expedite the persistence of regularity and continuous dependence of the data results, we shall recall[6] the notion of a *frequency envelope* (introduced in [25], [26] (see also [13], [19]) and refined slightly in [28] for applications to persistence of regularity).

Let M be an integer larger than s_0, and let $0 < \sigma \ll 1$ be a small number depending on M and s_0 to be chosen later. We allow our implicit constants in the notation $X \lesssim Y$ to depend on M and σ as well as s_0.

Definition 5.1. A *frequency envelope* c is a map from positive integers k to non-negative reals c_k such that

$$c_0 \sim 1, \tag{23}$$

$$\sum_{k=0}^{\infty} c_k^2 \lesssim 1 \tag{24}$$

and we have the log-Lipschitz conditions

$$2^{-Mr} c_k \le c_{k+r} \le 2^{\delta r} c_k \tag{25}$$

for all $k, r \ge 0$.

From (24) we see in particular that the c_k are bounded:

$$c_k \lesssim 1. \tag{26}$$

The extra condition (23) is a technical one which simplifies the treatment of the low frequencies (in particular, it allows one to exploit the L^2 conservation law easily). However, it is not strictly necessary and one can relax the notion of frequency envelope by dropping (23) (but then one has to estimate the low frequencies by iteration methods rather than via the conservation law). The condition (25) allows the coefficients c_k to drop quite fast as k increases, but only allows them to increase very slowly.

We now apply this concept of a frequency envelope to translation-invariant Banach spaces. If X is a translation-invariant Banach space, either of functions of space x or spacetime (t, x), and c is a frequency envelope, we define the envelope norm X_c by

$$\|f\|_{X_c} := \|f\|_X + \sup_{k \ge 0} \frac{\|P_k f\|_X}{c_k}. \tag{27}$$

It is easy to verify that X_c is indeed a norm. One can think of X_c as a Besov-type variant of X which contains more refined information about the distribution of the "energy" of f (as measured using the X norm) in frequency space.

---

[6]Our definition of frequency envelope has been modified for our subcritical setting, as opposed to the critical setting of the papers cited above. In particular, the frequency parameter k is restricted to non-negative values.

We now prove some linear and bilinear estimates involving envelope spaces that we will need in the next section. First observe that the Strichartz estimate (12) automatically implies the envelope variant

$$\|u\|_{S_c^k} \lesssim \|u(0)\|_{H_c^k} + \|u_t + Hu_{xx}\|_{(L_t^1 H_x^k)_c}; \tag{28}$$

this basically follows by applying (12) to the Littlewood-Paley components $P_j u$ of u and observing that the operator $\partial_t + H\partial_{xx}$ commutes with Littlewood-Paley operators. In a similar vein we see that the spaces S_c^k react well under differentiation, for instance we have

$$\|u\|_{S_c^2} \sim \|u_x\|_{S_c^1} + \|u\|_{S_c^1},$$

and we can omit the lower-order term $\|u\|_{S_c^1}$ if u is restricted to high frequencies $|\xi| \gtrsim 1$.

To prove some bilinear estimates, we shall need the following abstract lemma which allows us to convert bilinear estimates on Banach spaces X, Y, Z to bilinear estimates on their envelope counterparts X_c, Y_c, Z_c.

Let X, Y, Z be translation-invariant Banach spaces, and let $T : X \times Y \to Z$ be a bilinear operator. We say that T is of *product type* if it respects the Littlewood-Paley trichotomy; in other words, that $P_k T(P_{k_1} f, P_{k_2} g)$ vanishes unless we are in one of the following three cases:

- (Low-High interactions) $k = k_2 + O(1)$ and $k_1 \leq k + O(1)$.
- (High-Low interactions) $k = k_1 + O(1)$ and $k_2 \leq k + O(1)$.
- (High-High interactions) $k_1 \geq k + O(1)$ and $k_2 = k_1 + O(1)$.

It is easy to see that the product operator fg is of product type; it is clear that one can also add derivatives or frequency projections to the bilinear form fg and still retain product type, for instance $P_{+hi}(P_-(f_x)g)$ is of product type.

Lemma 5.2. *Let X, Y, Z be translation-invariant Banach spaces, and let c be a frequency envelope. Suppose that we have a bilinear operator T of product type which obeys the bilinear estimates*

$$\|T(f,g)\|_Z \lesssim \|f\|_X \|g\|_Y \tag{29}$$

for all test functions f, g. Suppose also that we have the "high-high refinement"

$$\|P_k T(P_{k_1} f, P_{k_2} g)\|_Z \lesssim 2^{-2\sigma(k_1 - k)} \|P_{k_1} f\|_X \|P_{k_2} g\|_Y \tag{30}$$

whenever $k_1 \geq k + O(1)$ and $k_2 = k_1 + O(1)$. Then we have

$$\|T(f,g)\|_{Z_c} \lesssim \|f\|_{X_c} \|g\|_{Y_c}.$$

Proof Since

$$\|T(f,g)\|_Z \lesssim \|f\|_X \|g\|_Y \lesssim \|f\|_{X_c} \|g\|_{Y_c}$$

it will suffice to prove the estimate

$$\|P_k T(f,g)\|_Z \lesssim c_k \|f\|_{X_c} \|g\|_{Y_c}.$$

We split f into Littlewood-Paley pieces $f = \sum_{k_1} P_{k_1} f$. First consider the contribution where $k_1 = k + O(1)$. Then by (29) we have

$$\|P_k T(P_{k_1} f, g)\|_Z \lesssim \|T(P_{k_1} f, g)\|_Z \lesssim \|P_{k_1} f\|_X \|g\|_Y \lesssim c_{k_1} \|f\|_{X_c} \|g\|_{Y_c}$$

which is acceptable since $c_{k_1} \sim c_k$ from (25).

Now consider the contributions where $k_1 < k - C$ for some large constant C. Then the Littlewood-Paley trichotomy allows us to restrict g to $\sum_{k_2 = k + O(1)} P_{k_2} g$, and we can estimate this contribution using the triangle inequality by

$$\sum_{k_2 = k+O(1)} \|P_k T(P_{<k-C} f, P_{k_2} g)\|_Z \lesssim \sup_{k_2 = k+O(1)} \|T(P_{<k-C} f, P_{k_2} g)\|_Z$$

$$\lesssim \|P_{<k-C} f\|_X \sup_{k_2 = k+O(1)} \|P_{k_2} g\|_Y$$

$$\lesssim \|f\|_{X_c} \sup_{k_2 = k+O(1)} c_{k_2} \|g\|_{Y_c}$$

which is acceptable as before.

Finally, consider the contributions where $k_1 > k + C$ for some large constant C. Then the Littlewood-Paley trichotomy allows us to restrict g to $\sum_{k_2 = k_1 + O(1)} P_{k_2} g$, and we can estimate this contribution using the triangle inequality by

$$\sum_{k_1 > k+C} \sum_{k_2 = k_1 + O(1)} \|P_k T(P_{k_1} f, P_{k_2} g)\|_Z.$$

Applying (30) we can estimate this by

$$\sum_{k_1 > k+C} \sum_{k_2 = k_1 + O(1)} 2^{-2\sigma(k_1 - k)} \|P_{k_1} f\|_X \|P_{k_2} g\|_Y,$$

which we can estimate by

$$\sum_{k_1 > k+C} \sum_{k_2 = k_1 + O(1)} 2^{-2\sigma(k_1 - k)} c_{k_1} \|f\|_{X_c} \|g\|_{Y_c}.$$

But this is acceptable by (25). ∎

Now we can state the bilinear estimates that we shall need.

Proposition 5.3. *For any functions f, g, ϕ on $\mathbf{R}$, we have*

$$\|fg\|_{H_c^j} \lesssim \|f\|_{H_c^2} \|g\|_{H_c^j} \tag{31}$$

and

$$\|f\phi\|_{H_c^j} \lesssim \|f\|_{H_c^j} (\|\phi\|_{L_x^\infty} + \|\phi_x\|_{H_x^{M+10}}) \tag{32}$$

for $k = 1, 2$. We also have the spacetime variants

$$\|fg\|_{(L_t^\infty H_x^j)_c} \lesssim \|f\|_{(L_t^\infty H_x^2)_c} \|g\|_{(L_t^\infty H_x^j)_c} \tag{33}$$

and

$$\|f\phi\|_{(L_t^\infty H_x^j)_c} \lesssim \|f\|_{(L_t^\infty H_x^j)_c} (\|\phi\|_{L_t^\infty L_x^\infty} + \|\phi_x\|_{L_t^\infty H_x^{M+10}}) \tag{34}$$

where now f, g, ϕ are functions on $[-1, 1] \times \mathbf{R}$. We also have the variant

$$\|f\phi\|_{S_c^1} \lesssim \|f\|_{S_c^1} (\|\phi\|_{L_t^\infty L_x^\infty} + \|\phi_x\|_{(L_t^\infty H_x^1)_c}). \tag{35}$$

Proof It suffices to prove the estimates (33), (34), (35), since the claims (31), (32) clearly follow as special cases.

We begin with (33). Using Lemma 5.2, it suffices to prove the non-envelope space estimate

$$\|fg\|_{L_t^\infty H_x^j} \lesssim \|f\|_{L_t^\infty H_x^2}\|g\|_{L_t^\infty H_x^j}$$

as well as the high-high refinement

$$\|P_k((P_{k_1}f)(P_{k_2}g))\|_{L_t^\infty H_x^j} \lesssim 2^{-2\sigma(k_1-k)}\|P_{k_1}f\|_{L_t^\infty H_x^2}\|P_{k_2}g\|_{L_t^\infty H_x^j}$$

whenever $k_1 = k_2 + O(1)$ and $k \leq k_1 + O(1)$.

The first inequality is clear from the Leibnitz, Hölder, and Sobolev inequalities, so we turn to the high-high refinement. For this we use Bernstein's inequality to estimate

$$\|P_k((P_{k_1}f)(P_{k_2}g))\|_{L_t^\infty H_x^j} \lesssim 2^{jk}2^{k/2}\|(P_{k_1}f)(P_{k_2}g)\|_{L_t^\infty L_x^1};$$

using Hölder we thus have

$$\|P_k((P_{k_1}f)(P_{k_2}g))\|_{L_t^\infty H_x^j} \lesssim 2^{jk}2^{k/2}2^{-2k_1}\|P_{k_1}f\|_{L_t^\infty H_x^2}2^{-jk_2}\|P_{k_2}g\|_{L_t^\infty H_x^j}$$

and the claim readily follows (in fact we have plenty of powers of 2^k and 2^{k_1} to spare).

Now we prove (34). A simple computation using Plancherel and (23), (25) shows that

$$\|\phi_x\|_{(L_t^\infty H_x^2)_c} \lesssim \|\phi_x\|_{L_t^\infty H_x^{M+10}}$$

so it suffices to prove the stronger estimate

$$\|f\phi\|_{(L_t^\infty H_x^j)_c} \lesssim \|f\|_{(L_t^\infty H_x^j)_c}(\|\phi\|_{L_t^\infty L_x^\infty} + \|\phi_x\|_{(L_t^\infty H_x^1)_c}). \tag{36}$$

We split $\phi = P_{lo}\phi + P_{hi}\phi$. Since $\|P_{hi}\phi\|_{(L_t^\infty H_x^2)_c} \lesssim \|\phi_x\|_{(L_t^\infty H_x^1)_c}$ so we see from (33) that the contribution of $P_{hi}\phi$ is acceptable. Now consider the contribution of $P_{lo}\phi$. We split $fP_{lo}\phi$ as $\sum_{k\geq 0}(P_k f)(P_{lo}\phi)$ and observe that the term $(P_k f)(P_{lo}\phi)$ has spatial Fourier support in the region $\{\xi : \langle\xi\rangle \sim 2^k\}$. Thus by (25) and dyadic decomposition it will suffice to prove the estimate

$$\|(P_k f)(P_{lo}\phi)\|_{L_t^\infty H_x^j} \lesssim \|P_k f\|_{L_t^\infty H_x^j}\|\phi\|_{L_t^\infty L_x^\infty}$$

for each $k \geq 0$. But we can replace the H_x^j norm on both sides by an L_x^2 norm, losing a factor of 2^{jk} on both sides, and the claim then follows from Hölder (noting that $\|P_{lo}\phi\|_{L_t^\infty L_x^\infty} \lesssim \|\phi\|_{L_t^\infty L_x^\infty}$). This proves (36) and hence (34).

Finally, we prove (35). In light of (36) it will suffice to show that

$$\|f\phi\|_{(L_t^4 C_x^1)_c} \lesssim \|f\|_{(L_t^4 C_x^1)_c}(\|\phi\|_{L_t^\infty L_x^\infty} + \|\phi_x\|_{(L_t^\infty H_x^1)_c}).$$

Again we split $\phi = P_{lo}\phi + P_{hi}\phi$. The treatment of $P_{lo}\phi$ is proven in exactly the same way as before. To estimate $P_{hi}\phi$, it suffices by arguing as before to prove the estimate

$$\|fg\|_{(L_t^4 C_x^1)_c} \lesssim \|f\|_{(L_t^4 C_x^1)_c}\|g\|_{(L_t^\infty H_x^2)_c}.$$

We again apply Lemma 5.2. The estimate

$$\|fg\|_{L_t^4 C_x^1} \lesssim \|f\|_{L_t^4 C_x^1}\|g\|_{L_t^\infty H_x^2}$$

can be easily deduced from Sobolev, Hölder, and the Leibnitz rule, so it suffices to prove the high-high refinement

$$\|P_k((P_{k_1}f)(P_{k_2}g))\|_{L_t^4 C_x^1} \lesssim 2^{-2\sigma(k_1-k)} \|P_{k_1}f\|_{L_t^4 C_x^1} \|P_{k_2}g\|_{L_t^\infty H_x^2}$$

whenever $k_1 = k_2 + O(1)$ and $k \le k_1 + O(1)$. But this is clear, for instance we may estimate

$$\begin{aligned}
\|P_k((P_{k_1}f)(P_{k_2}g))\|_{L_t^4 C_x^1} &\lesssim 2^k \|(P_{k_1}f)(P_{k_2}g)\|_{L_t^4 L_x^\infty} \\
&\lesssim 2^k \|P_{k_1}f\|_{L_t^4 L_x^\infty} \|P_{k_2}g\|_{L_t^\infty L_x^\infty} \\
&\lesssim 2^k 2^{-k_1} \|P_{k_1}f\|_{L_t^4 C_x^1} \|P_{k_2}g\|_{L_t^\infty H_x^2}
\end{aligned}$$

(in fact we can gain even more powers of 2^k and 2^{k_1} by being more careful). ∎

6. A FREQUENCY-ENVELOPE A PRIORI STRICHARTZ BOUND

We now prove a frequency envelope version of Theorem 3.1, namely:

Theorem 6.1. *Let c be a frequency envelope, and let u be an H_x^{M+10} solution to* (1) *with $u(0) = O_{H_c^1}(\varepsilon^2)$. Then we have*

$$u = O_{S_c^1([-1,1] \times \mathbf{R})}(\varepsilon^2).$$

Proof This will be a reprise[7] of the proof of Theorem 3.1. Again, we restrict all spacetime norms to the slab $[-1,1] \times \mathbf{R}$.

Fix c, u. We recall the fields $F, w, F_{LO}, F_{\pm HI}$ constructed in the proof of Theorem 3.1. To prove the desired bound on u, it will suffice by (14) to prove the estimates

$$\partial_x P_\pm F = O_{S_c^1}(\varepsilon^2); \quad w = O_{S_c^2}(\varepsilon^2). \tag{37}$$

By a standard continuity argument using the H_x^{M+10} global well-posedness theory (noting that the $L_t^\infty H_x^{M+10}$ norm will control the S_c^1 and S_c^2 norms thanks to (25)), it will suffice to prove this estimate assuming the bootstrap hypothesis

$$\partial_x P_\pm F = O_{S_c^1}(\varepsilon); \quad w = O_{S_c^2}(\varepsilon). \tag{38}$$

We now deduce (37) from (38). From (22), (23), and the low frequency projection we know that

$$\partial_x F_{LO} = O_{S_c^1}(\varepsilon^2) \tag{39}$$

so the low frequencies are fine.

Now we verify (37) for w. We begin with the initial data $w(0)$. From (14) we have

$$F_x(0) = O_{H_c^1}(\varepsilon^2) \tag{40}$$

so in particular we have

$$F_{HI}(0) = O_{H_c^2}(\varepsilon^2).$$

---

[7]In fact, this theorem supercedes Theorem 3.1, which is actually redundant in this argument; but we elected to place the non-envelope version of the theorem first in order to clarify the exposition.

From (31) and Taylor expansion we thus see (if ε is small) that

$$e^{-iF_{HI}(0)} = 1 + O_{H_c^2}(\varepsilon^2). \tag{41}$$

Now consider $F_{LO}(0)$. From (40) and frequency localization we have

$$\partial_x F_{LO}(0) = O_{H_x^{M+10}}(\varepsilon^2)$$

and hence by many applications of the chain rule (and using the boundedness of $e^{-iF_{LO}(0)}$)

$$\partial_x e^{-iF_{LO}(0)} = O_{H_x^{M+10}}(\varepsilon^2).$$

We can now multiply both sides of (41) by $e^{-iF_{LO}(0)}$ and use (32) to obtain

$$e^{-iF(0)} = e^{-iF_{LO}(0)} + O_{H_c^2}(\varepsilon^2).$$

Multiplying by $-iF_x(0)$ using (40), (31), (32) we obtain

$$\partial_x e^{-iF(0)} = -iF_x(0)e^{-iF(0)} = O_{H_c^1}(\varepsilon^2). \tag{42}$$

From (16) we thus have

$$w(0) = O_{H_c^2}(\varepsilon^2).$$

In order to verify (37) for w, it thus suffices by (28) to verify that

$$w_t + Hw_{xx} = O_{(L_t^1 H_x^2)_c}(\varepsilon^2).$$

By (17) it thus suffices to show that

$$P_{+hi}(P_-(F_{xx})w), P_{+hi}(P_-(F_{xx})P_{lo}(e^{-iF})) = O_{(L_t^1 H_x^2)_c}(\varepsilon^2).$$

The expression $P_{+hi}(P_-(F_{xx})P_{lo}(e^{-iF}))$ was already shown in the proof of Theorem 3.1 to be $O_{L_t^1 H_x^2}(\varepsilon^2)$, and is also of low frequency (supported on the region $|\xi| \leq 4$). Hence it is acceptable by (23). Thus it will suffice to consider the term $P_{+hi}(P_-(F_{xx})w)$. From (21) we see that $P_-(F_{xx}) = O_{(L_t^4 L_x^\infty)_c}(\varepsilon)$ and $w = O_{(L_t^\infty H_x^2)_c}(\varepsilon)$. So the claim will follow from the following paraproduct estimate.

Lemma 6.2. *For any functions $f(t,x)$ and $g(t,x)$, we have*

$$\|P_{+hi}(P_-(f)g)\|_{(L_t^1 H_x^2)_c} \lesssim \|P_-(f)\|_{(L_t^4 L_x^\infty)_c}\|g\|_{(L_t^\infty H_x^2)_c}.$$

Proof We use Lemma 5.2. From Lemma 3.2 and a Hölder in time we already have

$$\|P_{+hi}(P_-(f)g)\|_{L_t^1 H_x^2} \lesssim \|P_-(f)\|_{L_t^4 L_x^\infty}\|g\|_{L_t^\infty H_x^2}.$$

It thus suffices to show the high-high refinement

$$\|P_k P_{+hi}((P_-P_{k_1}f)P_{k_2}g)\|_{L_t^1 H_x^2} \lesssim 2^{-2\sigma(k_1-k)}\|P_-P_{k_1}(f)\|_{L_t^4 L_x^\infty}\|P_{k_2}g\|_{L_t^\infty H_x^2}$$

when $k_2 = k_1 + O(1)$ and $k \leq k_1 + O(1)$. But we can compute

$$\begin{aligned}
\|P_k P_{+hi}((P_-P_{k_1}f)P_{k_2}g)\|_{L_t^1 H_x^2} &\lesssim 2^{2k}\|(P_-P_{k_1}f)P_{k_2}g\|_{L_t^1 L_x^2} \\
&\lesssim 2^{2k}\|P_-P_{k_1}f\|_{L_t^4 L_x^\infty}\|P_{k_2}g\|_{L_t^\infty L_x^2} \\
&\lesssim 2^{2k}2^{-2k_2}\|P_-P_{k_1}f\|_{L_t^4 L_x^\infty}\|P_{k_2}g\|_{L_t^\infty H_x^2}
\end{aligned}$$

and the claim follows.   ■

Finally, we verify (37) for F_{HI}. By conjugation invariance of the S_c^k spaces it suffices to do this for F_{+HI}. By (18) it thus suffices to show that

$$e^{iF} w_x, e^{iF} E = O_{S^1}(\varepsilon^2).$$

From (21) we know that $F_x = O_{(L_t^\infty H_x^1)_c}(\varepsilon)$. By repeating the analysis used to prove (42) (using (33), (34) instead of (31), (32)), we thus have

$$\partial_x e^{\pm iF} = O_{(L_t^\infty H_x^1)_c}(\varepsilon). \tag{43}$$

By (35) we thus see that multiplication by $e^{\pm iF}$ is a bounded operation on S_c^1. Thus it will suffice to show that

$$w_x, E = O_{S_c^1}(\varepsilon^2).$$

The claim for w_x has already been proven, so it suffices to estimate E. We expand E using (19). Consider the first term $(P_{lo} + P_{-hi})(e^{-iF}(F_{+HI})_x)$ in (19). Note that we may freely replace e^{-iF} by $P_{-hi}e^{-iF}$ due to all the frequency projections. From (43) we have

$$P_{-hi}e^{-iF} = O_{(L_t^\infty H_x^2)_c}(\varepsilon).$$

In particular, from Sobolev we see that $P_{-hi}e^{-iF} = O_{L_t^\infty L_x^\infty}(\varepsilon)$. Since $(F_{+HI})_x = O_{S_c^1}(\varepsilon)$, we thus see from (35) that

$$(P_{lo} + P_{-hi})(e^{-iF}(F_{+HI})_x) = O_{S_c^1}(\varepsilon^2)$$

as desired.

The second term $P_{+hi}((F_{-HI})_x e^{-iF})$ of (19) is treated similarly, so we turn now to $P_{+hi}((F_{LO})_x e^{-iF})$. But this follows from (39) since multiplication by e^{-iF} is a bounded operation on S_c^1. This completes the proof of Theorem 6.1. ∎

As a particular corollary of Theorem 6.1, we see that if u is an H_x^{M+10} solution of (1) with $u(0) = O_{H_c^1}(\varepsilon^2)$, then we also have $u(t) = O_{H_c^1}(\varepsilon^2)$ for all $t \in [-1, 1]$, since the S_c^1 norm controls the $(L_t^\infty H_x^1)_c$ norm, which in turn controls the $L_t^\infty H_c^1$ norm. This can be thought of a strengthened version of the energy conservation law (9); it asserts not only that the H_x^1 norm stays bounded, but in fact its frequency profile is also stable for short times.

As an immediate consequence, we can prove (4) for H_x^{M+10} solutions. Indeed, if u is an H_x^{M+10} solution, we can define $a_0, a_1, \ldots$ to be the sequence $a_j := \|P_j u(0)\|_{H_x^1}$, and then let c_j be the sequence

$$c_j := 2^{-\sigma j} + \varepsilon^{-2}\Big(\sum_{0 \le k \le j} 2^{-M|k-j|} a_k + \sum_{k > j} 2^{-\sigma|k-j|} a_k \Big). \tag{44}$$

Then since $\|u(0)\|_{H_x^1} \lesssim \varepsilon^2$, it is easy to verify from Plancherel that c_j is a frequency envelope and

$$\|u(0)\|_{H_c^1} \lesssim \varepsilon^2.$$

Also, from (11) and Young's inequality we see that

$$\Big(\sum_j (2^{(s_0-1)j} c_j)^2 \Big)^{1/2} \lesssim 1 + \varepsilon^{-2} \Big(\sum_j (2^{(s_0-1)j} a_j)^2 \Big)^{1/2} \lesssim 1 + \varepsilon^{-2} \|u(0)\|_{H_x^{s_0}}. \tag{45}$$

Now from the corollary of Theorem 6.1 mentioned earlier, we have

$$\sup_{t\in[-1,1]} \|u(t)\|_{H_c^1} \lesssim \varepsilon^2.$$

From this, (45), and (11) we have

$$\sup_{t\in[-1,1]} \|u(t)\|_{H_x^{s_0}} \lesssim \|u(0)\|_{H_x^{s_0}} + \varepsilon^2$$

which is (4) for H_x^{M+10} solutions (note that by shrinking ε we can make ε^2 as small as $O(\|u(0)\|_{H_x^1})$, which is bounded by $O(\|u(0)\|_{H_x^{s_0}})$. In the next section we shall establish continuity in $H_x^{s_0}$, which will allow us to remove this a priori H_x^{M+10} restriction.

7. Conclusion of proof of Theorem 1.1

Now we improve the L_x^2 continuity to $H_x^{s_0}$ continuity for any $s_0 \geq 1$, using the corollary to Theorem 6.1. Let $u_0^{(n)}$ be any sequence of initial data in $H_x^{s_0} \cap B(0, R)$ which converges in $H_x^{s_0}$ norm to another initial datum u_0; we do not assume that u_0 is any smoother than $H_x^{s_0}$. Once again, we may normalize $R \ll \varepsilon^2$. We already know from Theorem 4.1 that $S(t)u_0^{(n)}$ converges to $S(t)u_0$ in the $C_t^0 L_x^2$ topology; we wish to improve this to $C_t^0 H_x^{s_0}$ convergence. Our basic tool here is the frequency envelope spaces $H_c^{s_0}$, which are in some sense compact in $H_x^{s_0}$.

Let $0 < \alpha$. We have to show that

$$\sup_{t\in[-1,1]} \|S(t)u_0^{(n)} - S(t)u_0\|_{H_x^{s_0}} \lesssim \alpha \tag{46}$$

for all sufficiently large n. Since $\|u_0\|_{H_x^{s_0}} < \infty$, we see from the monotone convergence theorem that there exists an integer $J \geq 1$ (depending[8] on u_0, α) such that

$$\|P_{>J} u_0\|_{H_x^{s_0}} \lesssim \alpha.$$

In particular, since $u_0^{(n)}$ converges to u_0 in $H_x^{s_0}$ we have

$$\|P_{>J} u_0^{(n)}\|_{H_x^{s_0}} \lesssim \alpha \tag{47}$$

for all sufficiently large n.

Fix n to be sufficiently large so that (47) holds. Let $a_0, a_1, \dots$ be the sequence (depending on n) defined by $a_j := \|P_j u_0^{(n)}\|_{H_x^1}$, and then let c_j be the sequence defined by (44). Then since $\|u_0^{(n)}\|_{H_x^1} \lesssim \varepsilon^2$, it is easy to verify that c_j is a frequency envelope and

$$\|u_0^{(n)}\|_{H_c^1} \lesssim \varepsilon^2.$$

---

[8]Note that this dependence on u_0 means that we establish merely that the solution map is continuous from $H_x^{s_0}$ to $H_x^{s_0}$; we know from the results of Koch-Tzvetkov [15] that the map cannot be *uniformly* continuous in this topology.

Also, from (47), (11) we see that

$$(\sum_{j \geq J+C} (2^{(s_0-1)j} c_j)^2)^{1/2} \lesssim \alpha/\varepsilon^2, \tag{48}$$

where the constant C depends on σ, α, ε but not on n. By the corollary[9] to Theorem 3.1 we thus have

$$\sup_{t \in [-1,1]} \|S(t)u_0^{(n)}\|_{H_c^1} \lesssim \varepsilon^2.$$

In particular, by (48) we have

$$\sup_{t \in [-1,1]} \|P_{>J} S(t)u_0^{(n)}\|_{H_x^{s_0}} \lesssim \alpha.$$

Taking limits as $n \to \infty$ and using Plancherel and Fatou's lemma, we obtain

$$\sup_{t \in [-1,1]} \|P_{>J} S(t)u_0\|_{H_x^{s_0}} \lesssim \alpha.$$

and hence

$$\sup_{t \in [-1,1]} \|P_{>J} (S(t)u_0^{(n)} - S(t)u_0)\|_{H_x^{s_0}} \lesssim \alpha.$$

Also, since $S(t)u_0^{(n)} - S(t)u_0$ converges to zero in $C_t^0 L_x^2$ norm from Theorem 4.1, we see that

$$\sup_{t \in [-1,1]} \|P_{\leq J} (S(t)u_0^{(n)} - S(t)u_0)\|_{H_x^{s_0}} \lesssim \alpha$$

for n sufficiently large (depending on α, J). The claim (46) then follows.

We have now shown that the solution map $S(t)$ can be uniquely continuously extended as a map from $H^{s_0} \cap B(0,R)$ (in the $H_x^{s_0}$ topology) to $H_x^{s_0}$; in particular we see by a standard limiting argument that the bound (4), which was already proven for $H_x^{M+10} \cap B(0,R)$ solutions, is in fact true for all $H^{s_0} \cap B(0,R)$ solutions.

Specializing to the case $s = 1$, we have established that $S(t)$ is a continuous map from $B(0,R)$ to H_x^1 in the H_x^1 topology. We can now easily take limits of (1) in the sense of distributions and conclude that (1) holds in the sense of distributions for all $u_0 \in B(0,R)$. This proves all the conclusions of Theorem 1.1, as desired.

8. Remarks

The methods employed here can certainly be extended below the H_x^1 regularity; we presented just the H_x^1 argument to simplify the exposition and to highlight the gauge transformation. Notice that the only estimate we used for the linear equation $u_t + H u_{xx} = 0$ were the Strichartz estimates. If instead one employed the Kato-type local smoothing estimate and the maximal function estimates as in [16], it is quite likely that one could lower the local existence result (or at least the a priori bound), probably to $H_x^{3/4+}$, in analogy with the situation for the KdV equation in [16]. The Lipschitz bound would then be more difficult to obtain, but perhaps one

---

[9]Strictly speaking, this was only proved for H_x^{M+10} solutions, but the claim follows for general solutions by a limiting argument using Fatou's lemma and the L_x^2 convergence result already proven.

can use the parametrices for the linearized equation $v_t + Hv_{xx} = (Uv)_x$ that were developed in [14]. Alternatively one can attempt to gauge away the worst part of the non-linearity $(Uv)_x$ by techniques similar[10] to that used for the a priori bound, e.g. by multiplying v by e^{-iF} where F is some sort of primitive of U as in (14). Of course, as one lowers the regularity $H_x^{s_0}$ in which one obtains well-posedness, it may well be necessary to lower the regularity used for the Lipschitz bound as well. Certainly one can only hope for a Lipschitz bound in a space no smoother than $H_x^{s_0-1/2}$, as the examples in [15] already show.

To go below $H_x^{3/4+}$, an obvious approach would be to introduce $X^{s,b}$ spaces, as used for instance in [3] or [17]. A preliminary analysis using the equation $w_t + iw_{xx} = P_{+hi}(P_-(\overline{w}_{xx})w)$ as a toy model[11] for (17) suggests that such an approach should be able to obtain local well-posedness all the way down to H_x^{0+}, and perhaps even to L_x^2 if one uses some sort of Besov refinement of the $X^{s,b}$ spaces (e.g. the space $X^{0,1/2,1}$; see [27]). However, there is a non-trivial difficulty in employing $X^{s,b}$ spaces or related norms, because such norms do not react well to gauge transformations such as $v \mapsto e^{-iF}v$. This is in contrast to Strichartz norms, which can easily absorb bounded factors such as e^{-iF} with little difficulty. It would be interesting to obtain a local well-posedness result at the L_x^2 regularity (which would immediately extend to a global well-posedness result, due to (6)) but we do not know how to achieve this. However it seems reasonably feasible to expect (for instance) a local well-posedness result at the level of $H_x^{1/2}$, which would then be global thanks to the conservation of the Hamiltonian (7). One may also be able to obtain global well-posedness at other regularities by use of such tools as the "I-method" (see e.g. [5]), although we do not pursue this question here.

The methods used here rely heavily on the fact that u is real (in order for the gauge transformation e^{-iF} to stay in L^∞). When u is complex we suspect that the Benjamin-Ono equation is in fact ill-posed even in smooth topologies, unless one imposes some sort of moment condition on u to eliminate the very low frequencies.

References

[1] M. Ablowitz, A. Fokas, *The inverse scattering transform for the Benjamin-Ono equation, a pivot for multidimensional problems*, Stud. Appl. Math. **68** (1983), 1–10.

[2] T. Benjamin, *Internal waves of permanent form in fluids of great depth*, J. Fluid Mech. **29** (1967), 559–562.

[3] J. Bourgain, *Fourier restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations*, Part I, Geometric and Funct. Anal. **3** (1993), 107-156.

[4] J.D. Cole, *On a quasi-linear parabolic equation occurring in aerodynamics*, Quart. Appl. Math. **9** (1951), 225-236.

[5] J. Colliander, M. Keel, G. Staffilani, T. Tao, H Takaoka, *Global well-posedness for the Schrodinger equations with derivative*, to appear, Siam J. Math.

[6] A.S. Fokas, L. Luo, *Global solutions and their asymptotic behavior for Benjamin-Ono-Burgers type equations*, Diff. Int. Eq. **13** (2000), 115–124.

---

[10]One possibility is that one would have to microlocalize the gauge transformation, applying a different gauge to each Littlewood-Paley piece, in the spirit of [25], [26], [13], [28].

[11]Of course, if we wish to prove well-posedness for u in H_x^s, then we expect to place w in H_x^{s+1}, so for this discussion one would try to iterate the toy model equation for $w(0)$ in H_x^{1+}.

[7] J. Ginibre, *Le probléme de Cauchy pour des EDP semi-linéaires périodiques en variables d'espace (d'aprés Bourgain)*, Séminaire Bourbaki 1994/1995, Asterisque **257** (1996), Exp. 796, 163–187.

[8] J. Ginibre, G. Velo, *Proprietes de lissage et existence de solutions pour l'equation de Benjamin-Ono generalisee*, C.R. Acad. Sci. Paris **308** (1989), 309–314.

[9] J. Ginibre, G. Velo, *Smoothing properties and existence of solutions for the generalized Benjamin-Ono equations*, J. Diff. Eq. **93** (1991), 150–232.

[10] E. Hopf, *The partial differential equation $u_t + uu_x = u_{xx}$*, Comm. Pure Appl. Math. **3** (1950), 201–230.

[11] R.J. Iorio Jr., *On the Cauchy problem for the Benjamin-Ono equation*, Comm. Partial Diff. Eq. **11** (1986), 1031–1181.

[12] M. Keel, T. Tao, *Endpoint Strichartz Estimates*, Amer. Math. J. 120 (1998), 955–980.

[13] S. Klainerman, I. Rodnianski, *On the global regularity of wave maps in the critical Sobolev norm*, IMRN **13** (2001), 656–677.

[14] H. Koch, N. Tzvetkov, *On the local well-posedness of the Benjamin-Ono equation in $H^s(\mathbf{R})$*, IMRN **26** (2003).

[15] H. Koch, N. Tzvetkov, *Nonlinear wave interactions for the Benjamin-Ono equation*, preprint.

[16] C. Kenig, G. Ponce, L. Vega, *Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle*, Comm. Pure Appl. Math. **46** (1993), 527–560.

[17] C. Kenig, G. Ponce, L. Vega, *Quadratic forms for the 1-D semilinear Schrodinger equation*, Trans. Amer. Math. Soc. **346** (1996), 3323–3353.

[18] S. Klainerman, M. Machedon, *Space-time Estimates for Null Forms and the Local Existence Theorem*, Comm. Pure Appl. Math., **46** (1993), 1221–1268.

[19] N. Masmoudi, K. Nakanishi, *From non-linear Klein-Gordon equation to a system of coupled Schrodinger equations*, Math Ann. **324** (2002) 2, 359–389.

[20] L. Molinet, J.-C. Saut, N. Tzvetkov, *Ill-posedness issues for the Benjamin-Ono and related equations*, SIAM J. Math. Anal. **33** (2001), 982–988.

[21] T. Muramatu, S. Taoka, *The initial value problem for the 1-D semilinear Schrodinger equation in Besov spaces*, preprint.

[22] K. Nakanishi, H. Takaoka, Y. Tsutsumi, *Counterexamples to bilinear estimates related with the KdV equation and the nonlinear Schrodinger equation*, Methods and Applications of Analysis **8** (2001), 569–578.

[23] G. Ponce, *On the global well-posedness of the Benjamin-Ono equation*, Diff. Int. Eq. **4** (1991), 527–542.

[24] J.C. Saut, *Sur quelques generalizations de l'equation de Korteweg-de Vries*, J. Math. Pures Appl. **58** (1979), 21-61.

[25] T. Tao, *Global regularity of wave maps I. Small critical Sobolev norm in high dimension*, IMRN **7** (2001), 299–328.

[26] T. Tao, *Global regularity of wave maps II. Small energy in two dimensions*, Comm. Math. Phys. **224** (2001), 443–544.

[27] D. Tataru, *On global existence and scattering for the wave maps equation*, to appear, Amer. J. Math.

[28] D. Tataru, *Rough solutions for the wave maps equation*, preprint.

[29] M.M. Tom, *Smoothing properties of some weak solutions of the Benjamin-Ono equation*, Diff. Int. Eq. **3** (1990) 683–694.

Department of Mathematics, UCLA, Los Angeles CA 90095-1555

E-mail address: tao@math.ucla.edu